PCT/SE 0 3 / 0 1 4 3 5

PRV

PATENT- OCH REGISTRERINGSVERKET
Patentavdelningen

REC'D 26 SEP 2003
WIPO PCT

Intyg
Certificate

Härmed intygas att bifogade kopior överensstämmer med de handlingar som ursprungligen ingivits till Patent- och registreringsverket i nedannämnda ansökan.

This is to certify that the annexed is a true copy of the documents as originally filed with the Patent- and Registration Office in connection with the following patent application.

(71) *Sökande* / *Applicant (s)* — *Amersham Biosciences AB, Uppsala SE*

(21) *Patentansökningsnummer* / *Patent application number* — *0203226-6*

(86) *Ingivningsdatum* / *Date of filing* — *2002-10-31*

Stockholm, 2003-09-18

För Patent- och registreringsverket
For the Patent- and Registration Office

Lisa Junegren

Avgift
Fee

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

PATENT- OCH REGISTRERINGSVERKET
SWEDEN

Postadress/Adress
Box 5055
S-102 42 STOCKHOLM

Telefon/Phone
+46 8 782 25 00
Vx 08-782 25 00

Telex
17978
PATOREG S

Telefax
+46 8 666 02 86
08-666 02 8

IMMUNOGLOBULIN BINDING SITE

Technical field

The present invention relates to the field of biochemistry, and more specifically to the identification of a novel binding site, which is strictly conserved among human IgGs of κ-type. The present invention also encompasses use of the novel binding site for identification and/or design of chemical entities capable of specific binding to such antibodies.

Background

The field of biochemistry began about a hundred years ago with a realisation that life processes involved phenomena that could be explained by the exact sciences of chemistry and physics. The early discoveries were mostly of general nature, but with time, the discipline of biochemistry matured and eventually became a well-accepted field as such. During the past decades, the growth within the field of biochemistry has been extensive and expansive, and numerous areas thereof are these days recognised, such as bioenergetics, molecular biology, membrane biochemistry, protein biochemistry, analytical biochemistry and many others.

A number of these areas utilise a steadily increasing number of biotechnological applications that involve antibodies, also known as immunoglobulins. As is well known, there are five different types of antibodies, namely immunoglobulin G (IgG), which is the most prevalent; immunoglobulin A (IgA); immunoglobulin M (IgM); immunoglobulin D (IgD); and immunoglobulin E (IgE). As is also well-known, antibodies can be prepared in two different forms, either as polyclonal antibodies, including various forms of antibodies, or as monoclonal antibodies, which is a form of pure antibodies produced by hybridomas. For many applications, the monoclonal antibodies are preferred. Examples of biotechnological applications of monoclonal antibodies are various immunochemical techniques, such as immunoaffinity extraction and chromatography,

immunochemical detectors, immunoblotting, receptor assays, enzyme inhibition assays, displacement assays and flow-injection immunoassays. For most medical applications, such as diagnosis, prevention and cure of disease, monoclonal antibodies are also most preferred, for example as biopharmaceuticals. At present, about thirty percent of the biotechnology-derived drugs under development are based on monoclonal antibodies of type G.

The Y-shaped disposition of the structure of the IgG molecule is well known from standard biochemistry textbooks. It is also well known that regarding its tertiary structure, one intact IgG molecule consists of six globular regions, each of which is formed by two domains. All domains in an IgG molecule have in turn similar structures, a characteristic fold, which has become known as the immunoblobulin fold. The secondary structure of this fold consists mainly of two beta sheets packed against each other. On the other hand, regarding its primary structure, IgGs consists of two light chains and two heavy chains, which are covalently, linked by four disulphide bridges strategically placed around the central juncture of the intact molecule also called the hinge region. The two globular parts, which correspond to the "base of the Y", form the Fc fragment and are formed by domains consisting of only heavy chain residues. Contrary to this, each of the "arms of the Y" constitute a Fab fragment with two globular parts each. Each of the globular parts in a Fab fragment is formed when one domain from the light chain contacts one domain from the heavy chain. It is well known that the globular part located further away from the centre of the antibody contains the so-called hypervariable regions and the antigen-binding site. The domains forming this part are known as V_L for the light chain domain and V_H for the heavy chain domain. On the other hand the globular part of the Fab fragment closer to the hinge region is formed by the so-called first constant domain of the heavy chain (CH1) and the constant domain of the light chain (CL). Correspondingly, the two globular parts forming the Fc fragment are formed one by two second constant domains (CH2) and the other by two third constant domains of the heavy chain (CH3).

By sequence homology, heavy chains of IgGs can be classified into four types 1,2,3 and 4 whereas light chains fall into two types called λ and κ. It is also well known that in humans about 40% of the IgG molecules carry a light chain of λ type whereas about 60% carry a light chain of κ type. IgGs which are built up of both light and heavy chains inherit both types of partitionings. Accordingly, one partitioning divides IgGs into four subclasses IgG1, IgG2, IgG3 and IgG4 as compared to the second partitioning which divides IgGs into two subtypes λ and κ. The same type of classification can be applied to antibody fragments like Fab fragments and so called $F(ab')_2$ fragments, which consist of two Fab fragments connected by a disulphide.

IgGs can be generated according to standard techniques in large quantities in cellular expression systems. The most widely used production method today includes purification via affinity chromatography based on the use of highly specific domains of proteins as affinity ligands. Illustrative examples of such IgG-binding protein ligands are protein A and protein G, which are cell wall proteins of the bacteria *Staphylococcus aureus* and group G *Streptococcus*, respectively. They both bind with different affinities to Fab and Fc fragments of various IgG types.

More specifically, protein A binds to IgG molecules from various mammals, with the highest affinity to the human subclasses of IgG1, IgG2 and IgG4. It binds primarily to a surface formed at the juncture of both the second and the third constant domains (CH2 and CH3) of IgG located on the Fc fragment, and can consequently not be used in affinity purification of other fragments of IgG such as Fab and so called $F(ab')_2$ fragments. Protein A binds to some Fab fragments however this binding is not generic since it targets the variable region. This lack of generality is a drawback under some circumstances, since the use of Fab and $F(ab')_2$ fragments has increased lately due to their considerably smaller size, as compared to intact IgG molecules, while still containing the functional antigen-binding region. For instance, the smaller size is an advan-

tage in the penetration of tumours with limited vascular supply in order to deliver cytotoxic payloads such as radionuclides, toxins, and chemotherapeutic agents to target cancer cells. On the other hand, protein G binds also to both Fc and Fab. Protein G binds partly to the same Fc fragment surface as protein A, but their ways of binding have been shown to be completely different. Protein G binds also to a highly conserved region of the constant part of the Fab fragment, primarily to residues from the heavy chain, and consequently it has potential to be used as a generic Fab binder. However, it has been reported that protein G has a reduced binding to Fab fragments of type IgG2. In addition to the above, protein ligands of this kind are often relatively expensive to produce, they are amenable to proteolytic degradation and they are also usually sensitive to both high and low pH values.

Accordingly, the development of novel and alternative ligands to IgG, which do not necessarily need to be proteins or even protein-based, is motivated. Such development would gain from a more thorough understanding of the binding properties of the IgG molecule. Even though methods for identification of novel ligands can be based on an experimental identification on a random basis, such as in screening, they still require use of a selected binding site or at least area on the target molecule. Moreover an alternative and in many cases complementary approach known from drug-discovery contexts as rational design requires knowledge of the three-dimensional structure of a limited region which can serve as a binding site.

In the random or screening approach, various methods have been suggested in the art for identification of chemical entities that bind specifically to an antibody or any target molecule in general. For example, phage display has been used to identify peptides with affinity to the Fc fragment of IgG. However, phage display can only produce peptidic ligands, which suffer from the above-discussed drawbacks related to degradation. Also, there is no guarantee as to

the generality of the binding since there is commonly no knowledge as to where on the target molecule used in screening the ligand actually binds.

Recently, computational tools have found a relatively widespread use in the field of understanding the binding properties of target molecules based on their known 3D structures. This is not in itself a new field. The pioneering work of B. Lee and F. M. Richards in the 1970s has inspired many investigators. However, as structural data and faster computers become available it is also less complicated to reveal the complex architecture of protein surfaces. For example, surfaces on proteins capable of interacting with binding molecules may include pockets, tunnels, channels, clefts and depressions. All this concepts may be covered by the more general term *cavity* the shape and accessibility of which, determines which concept is more appropriate. Accordingly, a depression is more accessible and flatter in shape than a *cleft*, which in turn is more accessible and flatter than a channel. Pockets and tunnels may be considered types of channels and as compared to a pocket a tunnel is more accessible since it has at least two entrances or connections to the outside of the protein. An additional type of cavity is the void. However a void is completely surrounded by protein atoms and therefore not accessible and therefore not appropriate for ligand binding.

Virtual methods have for example been suggested for identifying molecules capable of binding to proteins, and usually involve the identification of a cleft in the three dimensional structure of the target. However, if stronger binding is desired, a more advantageous conformation of the binding site can e.g. be a more pocket-like conformation, which spatially encloses the binding molecule to a larger extent than a cleft thereon allowing a maximisation of the number of possible interactions at the atomic level like hydrogen bonds, van der Waals, and electrostatic interactions, and therefore of the binding strength. This is especially the case if the ligand is a small organic molecule, which can have the advantage of being generally more stable than larger entities like peptides and

proteins. As compared to pockets and clefts, the depression is the less advantageous for binding a small molecule because it offers the smallest possibilities to complement the surface of the small molecule.

Thus, WO 01/37194 (Vertex Pharmaceuticals) discloses molecules and molecular complexes that comprise the active site binding pocket of the enzyme caspase-7. Methods are also disclosed, wherein the structural coordinates of caspase-7 are used to screen and design chemical entities that bind caspase-7 or homologues thereof. However finding a conserved binding pocket on the surface of an antibody is a more challenging task as compared to enzymes, which generally contain a substrate pocket related to their function. Antibodies like any other proteins that are not enzymes may or may not contain pockets or any other type of cavities. This is especially true for the constant domains of antibodies, which as compared to variable domains do not include the antigen-binding region that is often associated with a pocket.

Accordingly, despite the attempts that have been made so far, there is still a need in this field of alternative fast, easy and preferably easily applicable methods that are useful for the identification of novel IgG-binding molecules. As discussed above, since a useful starting point in the design of such novel methods is to select an advantageous binding site on the target molecule, there is also a need in this field of identifying novel binding sites within the IgG molecule. Moreover, if due to stability problems or other problems, it is required that the novel binding molecule is small, as compared to peptides and proteins that can be considered large in this context, then the novel binding site should be an accessible cavity. More specifically, in the order of decreasing preference the cavity should be a pocket, a tunnel or a cleft.

Summary of the present invention

Thus, the object of the present invention is to fulfil one or more of the above-discussed needs. More specifically, one object of the present invention is to

provide a tool useful e.g. for identification of novel IgG binding chemical entities and for other purposes. This can be achieved by the compound or binding pocket as defined in the appended claims.

A specific object of the invention is to provide a binding site in the constant region of the Fab part of an antibody, which binding site e.g. is suitable for use in the design of an antibody-selective medium. An even more specific object of the invention is to identify such a binding site in a region of the constant part of Fab, wherein the variability is as small as possible. Thus, yet another object is to provide a compound that can also be used for purification of any of the following: Fab fragments, $F(ab')_2$ fragments and intact IgGs of κ-type, or compositions that comprise one or more of the ones mentioned.

A further object of the present invention is to provide a method for identification of IgG ligands. This can be achieved by use of the compound or binding pocket according to the invention as defined in the appended claims. A specific object of the invention is to provide a method for virtual screening of small molecule ligands that are capable of binding to IgG.

Yet another object of the invention is to provide further uses of the compound or binding pocket according to the invention.

Further objects and advantages of the present invention will appear from the detailed description of the invention and the experimental part that follows.

Definitions

The term "binding pocket", as used herein, refers to a region of a molecule or molecular complex, that, as a result of its shape, favourably associates with another chemical entity.

The term "composite binding pocket" as used herein means a three-dimensional structure that is formed as a pocket between a light chain and a heavy chain of an antibody.

The term "interacting surface" means herein a surface comprised of residues capable of interacting with a binding molecule or other entity, e.g. by ionic attraction, hydrogen bonds, Van der Waals interaction etc.

The term "strictly conserved" is used herein to mean that after a sequence alignment of all sequences available from an internationally recognised sequence database (for instance the non-redundant database provided by the National Center for Biotechnology Information), the residue type is exactly the same at a specific position for all aligned sequences.

The terms "antibody of κ type", "Fab fragment of κ type" and "$F(ab')_2$ fragment of κ type" mean herein an antibody, a Fab fragment and an $F(ab')_2$ fragment respectively, wherein the light chain is of κ type.

The term "functional derivative" is used to mean a chemical substance that is related structurally and functionally to another substance. Thus, a functional derivative comprises a modified structure from the other substance, and maintains the function of the other substance, which in this instance means that it maintains the ability to interact with the same ligands. Thus, a "functional derivative" can be either a natural variation or fragment thereof, or a recombinantly produced entity. In addition, a "functional derivative" can also comprise added molecules or parts, as long as the described function is essentially retained.

The term "human κ-Fab constant part- comprising composition" means herein any composition comprising the globular region of an IgG molecule formed by the first constant domain of the heavy chain (CH1) and the constant domain of the light chain (CL). Thus the term includes any of the following terms which are well known from standard IgG terminology: Intact IgG molecules, $F(ab')_2$ fragments, Fab' fragments, Fab fragments and by definition the globular region named itself, all of which have human sequences and light chains of κ-type. This definition includes also any modifications of named IgG or named antibody fragments including even chimeric molecules formed in one part of one of said compositions and in another part of any of the following proteins, peptides, carbohydrates, lipids or any other organic or inorganic entity and chi-

meric combinations thereof and also any of the above-mentioned covalently attached to solid phase.

The term "structure coordinates" refers to Cartesian coordinates derived from for example mathematical equations related to the patterns obtained on diffraction of a monochromatic beam of X-rays by the atoms (scattering centres) of a protein or protein-ligand complex in crystal form. The diffraction data are used to calculate an electron density map of the repeating unit of the crystal. The electron density maps are then used to establish the positions of the individual atoms of the protein or protein complex.

The term "root mean square deviation" means the square root of the arithmetic mean of the squares of the deviations from the mean. It is a way to express the deviation or variation from a trend or object.

The term "docking" means herein a fitting operation, wherein the ability of a chemical entity to bind or "dock" to a binding site is evaluated.

The term "associating with" refers to a condition of proximity between a chemical entity, or portions thereof, and a target i.e. a binding pocket or binding site on a protein. The association may be non-covalent, wherein the juxtaposition is energetically favoured by hydrogen bonding or van der Waals or electrostatic interactions, or alternatively it may be covalent.

The term "library" means a collection of molecules or other chemical entities with different chemical structures and/or properties.

The term "query" means herein the definition of the criteria or properties of desired chemical entities that must be fulfilled for said entity to qualify as a hit in docking or screening. Accordingly, a "hit" means a chemical entity that fulfils the criteria of a query.

The term "chemical entity" is used herein for any molecule, chemical compound or complex of at least two chemical compounds and fragments of such compounds or complexes.

The term "ligand" means herein a chemical entity capable of specific binding to a target.

To "experimentally" contact a chemical entity with a target, or to "experimentally" provide a chemical entity, ligand or the like, means herein that it is provided physically, as opposed to virtually.

Brief description of the drawings

Figure 1 shows the structure coordinates of a human IgG of κ-type in the order. More specifically, Figure 1(a) shows the light chain of a human IgG of κ-type, while Figure 1(b) shows the structure coordinates of the heavy chain of a human IgG.

Figure 2 (a) and (b) show the alignment of human IgG Fab constant part light and heavy chain sequences of κ-type used in the identification of the binding pocket according to the invention.

Figure 3 shows an example of a query useful in 3D-*verify* mode in a docking step as used in example 2 below, wherein the residues belonging to the compound or binding pocket are shown in ball and stick.

Detailed description of the invention

The present invention provides a model for the three-dimensional structure of a novel binding pocket based on the structure coordinates of a human IgG of κ-type, as shown in Fig 1.

Thus, a first aspect of the present invention is an isolated compound, said compound having the structure of a human IgG binding pocket and comprising a first interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG κ light chain for the amino acids Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, and L181, and a second interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids P128, S129, L133, L150, K152, F175, P176, V178, L179, Q180, L184, L187 and S188, or a functional derivative of said compound. The above-defined amino acids are strictly conserved among human IgGs of κ-type. In the preferred embodiment, the present compound is limited to the part of an antibody that is re-

sponsible for the shape of the actual binding pocket and does not include the rest of a human IgG of κ type.

In one embodiment, the present invention is a binding pocket having the structure of a human IgG binding pocket and comprising a first interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG κ light chain for the amino acids Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, and L181, and a second interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids P128, S129, L133, L150, K152, F175, P176, V178, L179, Q180, L184, L187 and S188, or a functional derivative of said binding pocket. As mentioned above, the above-defined amino acids are strictly conserved among human IgGs of κ-type. In the preferred embodiment, the present binding pocket is an isolated binding pocket.

A specific embodiment of the present invention is a compound or binding pocket as defined above, wherein the second interacting surface is further defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids K126, F131, D153, S181, S182, and S186. Said amino acids are highly conserved between human IgGs of κ-type. In a specific embodiment, the highly conserved are conserved to at least about 85 %, in a more preferred embodiment they are conserved to at least about 90 % and in the most preferred embodiment, they are conserved to at least about 95 %.

In one embodiment of the present compound or binding pocket, the functional derivative thereof has a root mean square deviation from the backbone atoms of the binding pocket amino acids of not more than 2.0Å . In a preferred embodiment, said deviation is not more than about 1.5Å and in the most preferred embodiment, said deviation is not more than 1.0Å. The crystal structure presented herein, which was also used in the identification of the compound or binding pocket according to the invention, can also be found using the Protein Data

Bank accession code 1vge, Chacko et al., 1996. As the skilled in this field will realise, such structure coordinates usually exhibit some degree of variation due to e.g. thermal motion and slight differences in crystal packing as well as errors due to uncertainty arising from the finite resolution of the diffraction data and other errors, and compounds or binding pockets including such variations are accordingly encompassed by the present invention.

The binding pocket discussed above can also be denoted a "composite" binding pocket, since it is composed of two moieties originating from two different chains of an antibody, and more specifically from the defined, conserved amino acids of the constant 1 region of the heavy chain (CH1) of human IgG and the defined, conserved amino acids of the constant region of the κ type light chain of human IgG. Thus, in the native form of the antibody, the present binding pocket is formed between said two locations. In this context, it is to be understood that the term IgG includes herein all the human IgG sub-classes IgG1, IgG2, IgG3 and IgG4. Thus, the present invention discloses for the first time the above-defined binding pocket as isolated from its natural environment and useful as a general target for all or at least essentially all human κ-Fab constant part- comprising compositions.

An additional aspect of the present invention is a compound or a binding pocket, which corresponds to the interacting surfaces defined by the structure coordinates given herein, which are conserved among one or more of IgA, IgM, IgE and/or IgD, and which interacting surfaces define the shape of a pocket. Accordingly, such a compound or binding pocket is useful for the same applications as discussed herein, in order to define other group(s) of immunoglobulins, or subgroups thereof. Like the above-discussed IgG-binding pocket, the compound or binding pocket according to this aspect can be used to identify one or more of the other immunoglobulins and/or a composition that comprises the relevant part thereof.

The procedure by which the present inventors identified the present binding pocket will be described in detail in the experimental part below. In brief, previously, in research directed to identification of highly conserved regions of antibodies, the general approach has been to align both the constant regions of the human light and heavy chains. Contrary, the present inventors focussed on alignment of the constant region of human IgG κ light chains only, which quite unexpectedly enabled the identification of a strictly conserved region located between a light and a heavy chain of a human IgG κ antibody. No such high conservation has been reported in this exact region before where the conservation has been associated to human antibodies of κ type or fragments thereof. Further, no conserved pocket shaped binding sites have been disclosed before in this region. Thus, in "Introduction to Protein Structure" (C. Brandén and J. Tooze Second edition, 1998, Garland Publishing, Inc NY), it is mentioned that the region between the constant domains of Fab exhibit a high degree of conservation, but it is also disclosed how tightly they are packed against each other, as compared to the variable regions. Consequently, this reference in fact suggests that there is very little space, and certainly no pocket shaped space, between the Fab constant regions. In summary, the present invention was unexpected both because of the high degree of conservation of the binding site and because of the pocket shape thereof.

It will be readily apparent to those of skill in the art that the numbering of amino acids in other disclosures of human IgGs of κ type may be different than that presented herein. However, corresponding amino acids in such sequences are easily identified by visual inspection or by using commercially available homology software.

The binding pocket according to the invention is preferably prepared by isolation from a native source, i.e. a human IgG of κ type. Such isolation is easily performed by the skilled person following standard procedures e.g. from a cell line or a plasma sample. More specifically, the entire domain, which constitutes

the constant half of the Fab fragment, is isolated. For example, as it is well known in the field of preparing Fab fragments for different purposes to obtain the whole Fab from the intact IgG one can use papain, and to separate the two domains of Fab one could for example use an appropriate protease or proteases.

In a specific embodiment, the two entities that originate from the light and heavy chain, respectively, can be combined into one entity in any suitable way, e.g. by mutation of amino acid residues at specific locations in order to provide further disulphide bridge(s) between the fragments, and it is understood that any such modifications are also encompassed within the scope of the present invention

In one embodiment, the above-described compound or binding pocket has been complexed to an organic molecule, in which complex the binding constant is at least 10^{-4}, preferably at least 10^{-6} and most preferably at least 10^{-8} M. Thus, illustrative intervals are e.g. 10^{-4} to 10^{-8} M, such as 10^{-4} to 10^{-6} or 10^{-6} to 10^{-8}. Methods of identifying chemical entities that are capable of complexing with the present compound or binding pocket will be discussed in more detail below in relation to the second aspect of the invention. In this context, it is to be understood that the present use of the term "complex" is not intended to encompass a native human IgG of κ-type. In other words, the term "organic molecule" should not be interpreted as the remaining parts of such an IgG. As the skilled person will realise, the structure coordinates defined above for the binding pocket refer to the state before any complexing has occurred. Likewise, it is also realised that depending on the nature of the organic molecule, said structure coordinates might be slightly different in the actual complex. Sometimes an induced fit may occur.

In a specific embodiment, the complex is comprised of the above-described compound or binding pocket and a detectable label coupled thereto. In one embodiment, the label is not coupled to the interacting surfaces of the binding

pocket, but so as to leave them free for subsequent interaction. More specifically, the compound or binding pocket can be labelled with any suitable detectable label as conventionally used in immunoassays, such as a fluorescent label, a luminescent label, a chemiluminiscent label, an enzyme label, a radioactive label, an absorbance label etc. Such labelled compounds or binding pockets are useful e.g. in various assays for detection of human κ-Fab constant part- comprising composition thereof, as will be discussed in more detail below. The labelling of organic compounds or binding pockets with detectable labels is easily performed by the skilled person on this field using well known methods and reagents.

A second aspect of the present invention is a method of identification of a ligand for selective binding of a human κ-Fab constant part-comprising composition, wherein a compound or binding pocket as defined above is used. Accordingly, the novel compound or binding pocket according to the invention will provide a valuable target in research aimed at ligand design and/or identification. In one embodiment, the human κ-Fab constant part-comprising composition is a human IgG or a fragment thereof.

Thus, one embodiment of the present method is a method for evaluating the potential or ability of a chemical entity to associate with a human κ-Fab constant part-comprising composition, which method comprises to provide a library of chemical entities and screening of said library for ability to associate to a compound or binding pocket according to the invention. In an advantageous embodiment, the method also includes a further step of testing a selection of the chemical entities that associate to the compound or binding pocket by contacting them with a human κ-Fab constant part-comprising composition and grading said entities according to affinity. The library preferably comprises a large number of chemical entities and is screened for chemical entities, i.e. ligands, that bind to the compound or binding pocket. The library used may be comprised of random chemical entities or, in an alternative embodiment, the

library is a combinatorial library. The chemical entities may be naturally occurring or synthetic proteins, peptides, lipids, carbohydrates and any chimeric combinations thereon or any other organic or inorganic entities. In the most advantageous embodiment, the chemical entities of the library are relatively small organic molecules. In this context, "small" refers to molecules of a molecular weight below e.g. 1000 Da, preferably below about 500 Da.

In another embodiment, the present method is a structure-based or rational design of ligands capable of binding to the present compound or binding pocket. The method utilises the structure coordinates, or structure coordinates defining a selected region, as templates for the synthesis of ligands with strong and specific binding properties. Structure-based design is a well-known technology and the skilled person can readily perform this embodiment.

Thus, the present invention can also be a virtual method, in all or in parts. Accordingly, in one embodiment, the invention is a method for evaluating the potential or ability of a chemical entity to associate with a human κ-Fab constant part-comprising composition, which method comprises a first step wherein computational means are employed to perform a fitting operation between the chemical entity and a compound or binding pocket according to the invention and a second step wherein the results of said fitting operation are analysed to quantify the association between the chemical entity and the compound or binding pocket.

In a more specific embodiment, the method is a method of identifying a potential ligand to a human κ-Fab constant part-comprising composition, which method comprises

(a) generating a three-dimensional structure of a compound or binding pocket as defined above;

(b) employing said three-dimensional structure to design a candidate ligand;

(c) providing said candidate ligand;

(d) contacting the candidate ligand with a human κ-Fab constant part-comprising composition comprising said compound or binding pocket to verify any binding; and, optionally,

(e) repeating steps (b)-(d).

In one embodiment, the human κ-Fab constant part-comprising composition is a human IgG or a fragment thereof.

In one embodiment, step (c) involves to provide a virtual structure of the designed ligand, which is virtually contacted in step (d) with a virtual structure of the binding pocket. In an alternative embodiment, step (c) involves to provide the candidate ligand experimentally, in which embodiment the contact of step (d) is also performed experimentally.

There are many commercial tools available for virtual methods of this kind. Examples of commercially available specialised computer programs that are useful in the process of selecting fragments or chemical entities are e.g. GRID (available from Oxford University, Oxford, UK); MCSS (available from Accelrys formerly MSI, San Diego); AUTODOCK (available from Scripps Research Institute, La Jolla); UNITY and FLEXX (available from Tripos Associates, St. Louis, Mo.) and DOCK (available from University of California, San Francisco).

Examples of software useful in connecting chemical entities or fragments include CAVEAT (available from University of California, Berkley), HOOK (available from Accelrys formerly MSI, San Diego); and 3D Database systems such as ISIS (MDL Information Systems, San Leandro).

Finally, programs for *de novo* ligand design methods include e.g. LUDI (available from Accelrys formerly MSI San Diego); LeapFrog (available from Tripos

Associates, St. Louis, Mo.) and SPROUT (available from the University of Leeds, UK).

Once a chemical entity has been designed or selected, the efficiency with which it binds to the binding pocket may be tested and optimised by computational evaluation. For example, a relatively small difference in energy between its free and bound states, i.e. small deformation energy of binding, is desired. Alternatively, ligands are prepared and tested in standard experiments in the lab.

In a specific embodiment, the method is a method for evaluating the potential or ability of a chemical entity to associate with a human κ-Fab constant part-comprising composition, which method comprises the steps of

(a) providing a virtual library of chemical entities;

(b) docking the chemical entities to a compound or a binding pocket as defined above;

(c) defining at least one query based on the results of the docking operation;

(d) screening all entities docked in step (b) while in the docked conformation with the query defined in step (c) for evaluating the potential or ability thereof to associate to the compound or binding pocket;

(e) inspection and, optionally, removal of redundancy; and

(f) providing one or more of the chemical entities that associated with the compound or binding pocket and experimentally testing their binding to a human κ-Fab constant part-comprising composition; and, if more than one chemical entity was tested,

(g) rating the affinities thereof to human κ-Fab constant part-comprising composition. In one embodiment, the human κ-Fab constant part-comprising composition is a human IgG or a fragment thereof.

For practical reasons, the virtual library used in the docking should comprise a limited number of chemical entities and it is preferable to reduce redundancy

among said entities. Accordingly, in one embodiment, the starting material is library that comprises an already diverse selection of entities suitable for use in the docking.

In an alternative embodiment, a virtual library is provided in step (a) that consists of one conformation of the 3D structure of chemical entities, which are either commercially available or are synthesised according to known methods. An example of such a library can be prepared by exporting a file containing information of the 2D structure of the chemical entities, which are normally provided by vendors of chemical entities. The 2D structures can then be used to produce one 3D conformation by using standard molecular modelling programs that are commercially available. One example of such a program is CONCORD available from Tripos Associates, St. Louis, Mo. Further, step (a) of this embodiment comprises a further step of filtering and removal of redundancy among the entities of the library provided. The removal may be a filtering that excludes entities in accordance with certain predetermined criteria. Examples of useful filtration criteria are molecular weight and the calculated water/octanol partition coefficient. In the present method, where the goal of the screening is to obtain a putative ligand to a rather small pocket which first will be tested in solution for binding to the target protein, a suitable range of molecular weight is 200-500 Da while the calculated water/octanol coefficient can be set to be lower than 4.0. Depending on the intended use of the ligand, additional or other criteria can be set.

Further, step (b) analyses the fit of each one of the chemical entities to the compound or binding pocket according to the invention. "Docking" means in this context the use of computational tools and available structural data to obtain new information about binding modes and molecular interactions. Thus, docking is the placement of a putative ligand in an appropriate configuration for interacting with a binding site. The database used for docking should contain a large number of diverse chemical entities, and it can be prepared for the

purpose or be obtained from commercial sources. Programs for faster however more information-requiring forms of docking are also commercially available, either for searching with fixed or flexible rotational bonds, such as the UNITY *3D search* algorithm, or the UNITY *flexible 3D search* algorithm available from Tripos Associates, St. Louis Mo. In general, docking can be accomplished by geometric matching of a ligand and its binding site, or by minimising the energy of interaction. As the skilled person in this field will know, geometric matching is faster, and searching with fixed rotational bonds is faster than with flexible ones. In the same way, docking may include flexibility of the side-chains or it may keep them fixed. In an advantageous embodiment, the results of the docking operation of step (b) are evaluated by visual inspection of the extent of contact between the interacting surface of the compound or binding pocket and the putative ligand. Additionally, or alternatively, the gaps formed between the two are calculated with the help of at least one query defined in step (c) and applied in step (d). Thus, step (d) is a query match screening wherein the coordinates of the screened entities are the docked conformation coordinates obtained from step (b) which are kept fixed during the screening procedure. Then, the positive hits from step (d) can be visually inspected in stereo-graphics to remove molecules that do not associate with the compound or binding pocket and/or to remove molecules that are similar to selected ones in order to further discriminate against redundancy in step (e). To still be considered as a hit, the chemical entity should be complementary to the compound or binding pocket with respect to conformation, hydrogen bonds, charge and/or hydrophobicity. Most preferably, all this aspects are satisfied.

For reasons of simplicity, in the most advantageous embodiment, step (e) is preferably a visual inspection. However, the present invention also encompasses alternative embodiments where the inspection is performed for example by computational means.

In step (f), one or more of the chemical entities that associated with the compound or binding pocket during the screening according to step (d) are selected as candidates for further testing, which preferably means that they are provided experimentally. Many chemical entities that are present in commercial databases are also commercially available and hence easily purchased. Alternatively, the candidates are easily synthesised in accordance with standard methods. In one embodiment, the binding experiments are simply to contact the candidate(s) with a human κ-Fab constant part-comprising composition in solution by means of for instance NMR and/or Surface Plasmon Resonance techniques to evaluate any complexing. As the skilled person in this field will appreciate, in order to analyse the exact location where the candidate interacts with the compound or binding pocket, more complex methods will be required, such as crystal structure determination of complexes or alternative NMR techniques. Accordingly, a specific embodiment of the present method is a method for evaluating the potential or ability of a chemical entity to associate with a human IgG, a Fab fragment or a $F(ab')_2$ fragment of κ-type by binding to the compound or binding pocket according to the invention. Methods for testing binding of a ligand to a binding site are well known in this field and hence easily performed by the skilled person in this field in accordance with routine experiments.

A third aspect of the invention is the use of a compound or binding pocket according to the invention for identification or design of a chemical entity capable of selective binding of a human κ-Fab constant part-comprising composition. In one embodiment, the human κ-Fab constant part-comprising composition is a human IgG or a fragment thereof. Some embodiments of this aspect have been described in detail above. Thus, the products resulting from the above-described methods, i.e. the selectively binding chemical entities, are useful e.g. as ligands in chromatography methods.

Another aspect is the use of a compound or binding pocket as defined above for site-specific modification of a human κ-Fab constant part-comprising composition. In one embodiment, the human κ-Fab constant part-comprising composition is a human IgG or a fragment thereof. More specifically, a human κ-Fab constant part-comprising composition can be modified by binding a suitable chemical entity selectively to the compound or binding pocket identified by the present inventors. In an advantageous embodiment, said modification is a stabilisation of Fab-folding by binding a ligand to the compound or binding pocket.

The compound or binding pocket defined above is also useful in assays wherein human κ-Fab constant part-comprising composition are detected, in which case it is preferably labelled with a suitable detectable label as discussed above. Such assays may be in solution or on solid phase. In one embodiment, the human κ-Fab constant part-comprising composition is a human IgG or a fragment thereof. In the preferred embodiment, the present assay is a competitive assay, wherein the ability of a candidate ligand to displace a known ligand's binding to a compound or binding pocket as defined above is evaluated.

In another embodiment, the compound or binding pocket defined above is used in an immunological assay for detection of a human κ-Fab constant part-comprising composition.

The compound or binding pocket may be used to bind cytotoxic molecules and compositions, such as radionuclides, toxins and chemoterapeutic agents that are released when the antibody associates with a cell that threatens the health of the host. More specifically, the target for the antibody can be an antigen located in for instance a cancer cell.

In addition, the present compound or binding pocket is also useful in various medical applications. Binding pockets, also referred to as binding sites in the

present invention, are of significant utility in fields such as drug discovery. The association of natural ligands with the binding pocket of an antibody may prove to be the basis of biological mechanisms of action.

The present invention also encompasses a computer for producing a three-dimensional representation of a compound or binding pocket according to the invention, which computer comprises

(i) a computer-readable data storage medium comprising a data storage material encoded with computer-readable data, wherein said data comprises the structure coordinates as shown in Figure 1 for an IgG κ light chain for the amino acids Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, and L181 and the structure coordinates as shown in Figure 1 for an IgG heavy chain for the amino acids P128, S129, L133, L150, K152, F175, P176, V178, L179, Q180, L184, L187 and S188;

(ii) a working memory for storing instructions for processing said computer-readable data;

(iii) a central-processing unit coupled to said working memory and to said computer-readable data storage medium for processing said computer-machine readable data into said three-dimensional representation; and

(iv) a display coupled to said central-processing unit for displaying said three-dimensional representation.

In a specific embodiment, the computer-readable data further comprises the structure coordinates as shown in Figure 1 for an IgG heavy chain for amino acids K126, F131, D153, S181, S182, and S186.

Thus, the computer is producing a three-dimensional graphical structure of a compound or binding pocket as defined above. Such a graphical structure is for example useful in the methods described above, wherein selectively binding entities are designed and/or identified.

The computer according to the invention comprises standard components, for example as discussed in more detail in US patent no. 6,183,12.

Another aspect of the invention is a machine-readable datastorage medium comprising a data storage material encoded with machine readable data, wherein said data is defined by all or a portion of the structure coordinates of a compound or binding pocket according to the invention. Such a datastorage medium is for example useful in the methods described above.

Detailed description of the drawings

Figure 1 (a) and (b) show the structure coordinates of the light chain and the heavy chain of, respectively, of a human IgG of κ-type. The structure coordinates given are for the amino acids identified as strictly conserved and highly conserved by the present inventors, and they are provided with the numbering conventionally used for the full sequence. The full amino acid sequence together with structure coordinate data for all amino acids can be found at for instance http://www.rcsb.org/pdb/. The following abbreviations are used in Fig. 1:
"Atom type" refers to the element whose coordinates are measured. The first letter in the column defines the element.
"X, Y, Z" crystallographically define the atomic position of the element measured.
"B" is a thermal factor that measures movement of the atom around its atomic centre.
"Occ" is an occupancy factor that refers to the fraction of the molecules in which each atom occupies the position specified by the coordinates. A value of "1" indicates that each atom has the same conformation, i.e., the same position, in all molecules of the crystal.

Figure 2 (a) and (b) show the alignment of human IgG Fab constant part light and heavy chain sequences of κ-type used in the identification of the binding pocket according to the invention.

Figure 3 shows an example of a query useful in 3D-*verify* mode in a docking step as used in example 2 below, wherein the residues belonging to the compound or binding pocket are shown in ball and stick. According to this query, to be considered a hit, a docked molecule should have five- or six-membered rings with centra located within each of the spheres. There are several hydrogen bond donors at the entrance of the cavity, for instance atoms from the side chains of Thr-180 and Gln-160 from the light chain and Ser-186 from the heavy chain. Possible hits might thus probably contribute with acceptors. In addition, the ligand should fit into the yellow surface of the binding site. A similar query (Query 2) was also defined by dropping the requirement of ring with centre inside the largest sphere located at the entrance of the pocket.

EXPERIMENTAL PART

The present examples are provided for illustrative purposes only and are not to be construed as limiting the present invention as defined by the appended claims. All references given below and elsewhere in the present specification are hereby incorporated by reference.

Example 1: Identification of the binding pocket according to the invention

To identify conserved sequence patches in the constant regions of heavy and light chain sequences of human Fab-fragments of κ-type sequence homology searches using BLAST (Altschul, 1990) followed by sequence alignments using CLUSTAL W (Thompson, 1994) were performed. A total of 29 heavy chain and nine κ light chain sequences of human IgG's were identified after the BLAST search. (Figure 2).

The highest-resolution (2.0 Å) crystal structure of κ-Fab was investigated (accession code to the Protein Data Bank 1vge, Chacko *et al.*, 1996). The MOLCAD (program available from Tripos Associates, St. Louis, Mo.) *multi channel surface* tool was used to identify possible binding sites in the constant part.

Two clefts and one pocket were identified. The pocket (Figure 3) is located between the constant parts of the light and the heavy chain. Strictly conserved or highly conserved residues surround this pocket. Because of this conservation and since a small molecule might have higher affinity towards an invagination than a more open binding site this pocket was chosen as target. The residues forming the pocket together with some residues located at the entrance and contributing significantly to the topology of the putative binding site were identified. From the light chain these are **Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, L181** and they are all strictly conserved for all sequences of κ-type aligned. The residues from the heavy chain are (bold strictly- and remaining highly conserved) K126, **P128, S129,** F131, **L133, L150, K152,** D153, **F175, P176, V178, L179, Q180,** S181, S182, **L184,** S186, **L187** and **S188.**

Example 2: Use of the binding pocket to identify selectively compounds

In the example below, the term "compound" is sometimes used to denote chemical entities tested for their ability to bind selectively to a binding pocket according to the invention. However, as appears clearly from the context, such "compounds" are not the claimed compounds discussed in the section "Detailed description of the invention" and in the appended claims.

The program package SYBYL version 6.7 (available from Tripos Associates, St. Louis, Mo.) running on an OCTANE 2-CPU 195MHz Silicon Graphics workstation was used for all modelling. This interface provides the necessary information regarding the software and databases below.

Virtual library

The program SELECTOR was used for filtering the molecules in MDL™ (MDL Information Systems Inc.) Available Chemical directory (ACD) allowing only for entities with a molecular weight in the range 200-500 Da and a calculated water/octanol partition coefficient (ClogP) less than 4.0. The limit

500 Da was used since it was assumed that the binding site was not appropriate to accommodate larger ligands. Also smaller entities with fewer degrees of freedom are more suitable for computational methods. Entities containing tri-phosphate or tri-peptide substructures were also rejected. After filtering the number of ACD molecules was reduced to about 111.000. A distance-based algorithm as implemented in the program Diverse Solutions version 4.04 (Pearlman, et al, 1998; available from Tripos Associates, St. Louis, Mo.) was used to select one diverse subset with 50.000 molecules for the docking database. According to Potter and Matter (Pötter & Matter, 1998 Pötter, T & Matter, H. (1998) Random or rational design? Evaluation of diverse compound subsets from chemical structure databases. J. Med. Chem. 41:478-488) a database may be considered to be optimally diverse if the mean Tanimoto to the nearest neighbour is 0.85 and the standard deviation is small. The docking database of 50.000 molecules had a mean Tanomoto of 0.81 and a standard deviation of 0.11. The 3D structures were generated with CONCORD version 4.04 (Pearlman, 2001 available from Tripos Associates, St. Louis, Mo.). Molecules in the docking database were also ionised to represent their protonation state at neutral pH and minimised in 500 cycles using the MMFF94 force field (Halgren, 1996; available from Tripos Associates, St. Louis, Mo.). To increase the docking performance (speed) they were divided in smaller sets of 500 each.

Docking using FlexX

Docking simulations were performed with FlexX The protein structure used was the structure named above and used for the identification of the pocket. In the protein structure, the ε carbonyl oxygen of H:Gln-180 is located 2.5Å away from one of the δ carboxyl oxygens of H:Asp153. This was assumed to be an error due to misinterpretation of the electron density of the carboxyamide terminal group of H:Gln-180, and consequently this group was flipped around 180°. In this corrected structure, the ε nitrogen of H:Gln-180 is at favourable hydrogen bonding distance to the carboxyl oxygen of H:Asp153. Otherwise, defaults have been used when creating the rd file and no special customisations

were made. The residues belonging to the active site file in the rd file are the same as those surrounding one binding pocket identified as described in example 1, and some residues located in the surroundings and contributing to the topology of the putative binding site Prior to docking, all water molecules were removed. The best ranked conformation and its FlexX score were saved for each molecule.

Defining queries with UNITY verify 3D search

A quick analysis of some thousand docked molecules inspired to the definition of two queries to extract the molecules, which actually docked inside the pocket. According to one of them (Query 1, Figure 3), a docked molecule should, to be considered as a hit have five- or six-member rings with centra located within each one of two spheres. The smallest sphere (radius 2.0 Å) is centred inside the pocket and the largest one (radius 3.0 Å) is centred at the entrance. In a second query (Query 2), the requirement of the ring at the entrance was dropped. All hits from Query 1 should also be hits from Query 2. It might be argued that Query 1 is not contributing with new hits, but for bookkeeping reasons it might be useful to know which molecules fulfil the more demanding Query 1.

Virtual screening of the docked molecules with UNITY in 3D verify mode

This step was performed using the two related queries defined in the previous section.

Criteria for hit extraction after visual inspection

To be considered a hit, the ligand should be complementary to the binding site with respect to conformation (shape), hydrogen bonds, charge and hydrophobicity. These criteria are based on statistical analysis of high-resolution protein structures which have shown that less than 2% of the polar atoms are buried without forming a hydrogen bond (McDonald, 1994 McDonald IK, Thornton JM. (1994), *J. Mol. Biol.*, 238, 777-793.), and the increase in entropy as hydro-

phobic surfaces meet and water molecules are released (Tanford, 1980 Tanford C, (1980) The Hydrophobic Effect, 2nd ed. Wilery, New York,). Complementary in shape should maximise the number of possible polar and hydrophobic interactions. In addition, the ligand should be as rigid as possible and bind in a low-energy conformation to reduce the total free energy of binding.

Virtual screening results

A total of 43031 entities docked to the binding site with a favourable (negative) estimated free energy of binding, of these 98 satisfied Query 1 and 151 the less demanding Query 2. The difference, 53 entities, satisfied Query 2 but not Query 1. After visual inspection, 58 from the first set and 26 from the last set were selected. These 84 were ordered from suppliers, 76 thereof were delivered and 46 thereof turned out to be soluble as required.

Screening using NMR

All NMR experiments were performed at 298 K on a Bruker Avance 500 MHz spectrometer. The 1D saturation transfer difference method (STD NMR) was used as screening assay (Mayer M. and Meyer B. 1999. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem., Int. Ed.* **38:** 1784-1788) using several antibody concentrations in order to differentiate between compounds with different binding strength (Peng J. W., Lepre C. A., Fejzo J., Abdul-Manan N. and Moore J. M. 2001. Nuclear Magnetic Resonance-Based Approaches for Lead Generation in Drug Discovery. Methods in Enzymology. **338:** 202-230). The antibody used was a human Fab of κ–type. In all cases ligands were tested one-by-one. On-resonance irradiation was set at 0 ppm and off-resonance irradiation was set at –40 ppm. Irradiation time in each scan was 2 s and 16K data points were collected with 1024 scans in total. Compounds for testing were dissolved in $DMSO_{d6}$ to a concentration of 50 mM and 5 µL of the concentrated ligand solution was added to 495 µL buffer solution. The samples thus consisted of 0.5 mM ligand, 20 mM

phosphate buffer, 100 mM NaCl and 5% $DMSO_{d6}$ in D_2O at pD 7.5, uncorrected reading on pH-meter.

Compounds were initially tested for binding with 0.5 μM antibody. Interesting ligands were further tested with protein concentrations of 100 or 20 nM. A one-dimensional ^1H-spectrum was acquired first and subsequently a saturation transfer difference (STD) spectrum was acquired. Each analysis took 60 minutes on the spectrometer. A positive result was obtained if signals from the ligand were observed in the difference spectrum. The analysis was setup for automation so that several samples could be analysed over night (usually 10-15 samples/night).

Results

Binding test

As many as 46 of the virtual screening hits were tested with NMR according to the procedure described above. The results from NMR screening together with additional compound data are compiled in Table 1 below. In total 24 compounds gave a positive result in the first round with the highest antibody concentration (500 nM). These 24 compounds were subjected to a second round with an antibody concentration of 100 nM. The 4 compounds showing the strongest signal in the second round were tested for binding in a third round with 20 nM antibody. Three of the compounds in the third round gave a positive result and where thus designated as the strongest binders to the antibody out of the 46 tested compounds.

Table 1. Results from the NMR screening

Clogp is the calculated water/octanol partition coefficient. Concentration code as follows: conc. 1 means 500, conc 2 100 and conc 3 20 nM antibody. NMR signal code: 0 no, 1 weak and 2 strong signal.

Compound no.	Clogp	Mw	conc 1	conc 2	conc 3
1	0,2	214	0		
2	2,0	209	0		
3	2,0	266	0		
4	3,6	326	1	0	
5	2,4	207	1	1	
6	2,1	259	0		
7	0,1	257	0		
8	2,2	260	2	0	
9	0,0	239	0		
10	-0,4	236	0		
11	2,2	234	0		
12	3,4	299	0		
13	2,0	301	0		
14	2,4	261	1	0	
15	1,6	253	0		
16	2,7	214	2	1	
17	2,4	252	2	2	1
18	2,3	272	0		
19	2,1	222	0		
20	2,1	243	2	2	0
21	1,9	208	0		
22	2,8	218	0		
23	3,6	219	1	0	
24	2,1	245	2	2	2
25	3,8	278	1	0	
26	3,2	209	1	0	
27	3,4	228	1	1	
28	3,8	268	2	0	
29	2,3	281	1	1	
30	1,8	242	0		
31	2,1	208	1	0	
32	0,9	217	0		
33	1,9	278	0		

34	2,1	243	2	0	
35	5,0	288	0		
36	-0,4	216	0		
37	2,8	249	2	0	
38	3,5	202	1	0	
39	2,8	212	2	0	
40	2,3	299	0		
41	2,7	243	1	1	
42	2,8	247	2	2	1
43	2,8	356	0		
44	4,1	308	1	0	
45	3,2	334	2	1	
46	1,2	256	1	0	

CLAIMS

1. An isolated compound, said compound having the structure of a human IgG binding pocket and comprising a first interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG κ light chain for the amino acids Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, and L181, and a second interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids P128, S129, L133, L150, K152, F175, P176, V178, L179, Q180, L184, L187 and S188, or a functional derivative of said compound.
2. A compound according to claim 1, wherein the second interacting surface is further defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids K126, F131, D153, S181, S182, and S186.
3. A compound according to claim 1 or 2, wherein the functional derivative of the compound has a root mean square deviation from the backbone atoms of the binding pocket amino acids of not more than 2.0Å.
4. A binding pocket having the structure of a human IgG binding pocket and comprising a first interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG κ light chain for the amino acids Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, and L181, and a second interacting surface, which is defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids P128, S129, L133, L150, K152, F175, P176, V178, L179, Q180, L184, L187 and S188, or a functional derivative of said binding pocket.
5. A binding pocket according to claim 4, wherein the second interacting surface is further defined by the structure coordinates shown in Fig 1 for an IgG heavy chain for the amino acids K126, F131, D153, S181, S182, and S186.

6. A compound according to claim 4 or 5, wherein the functional derivative of the binding pocket has a root mean square deviation from the backbone atoms of said amino acids of not more than 2.0Å.
7. A complex comprising a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6 coupled to an organic molecule, in which complex the binding constant is at least 10^{-4} M.
8. A complex according to claim 7, which is comprised of said compound or binding pocket and a detectable label coupled thereto.
9. A method for evaluating the potential or ability of a chemical entity to associate with a human κ-Fab constant part-comprising composition, which method comprises to provide a library of chemical entities and screening said library for ability to associate to a compound according to any one of claims 1-3 or to a binding pocket according to any one of claims 4-6.
10. A method according to claim 9, which includes a further step of testing a selection of the chemical entities that associate to said compound or binding pocket by contacting them with a human κ-Fab constant part-comprising composition and grading said entities according to affinity.
11. A method for evaluating the potential or ability of a chemical entity to associate with a human κ-Fab constant part-comprising composition, which method comprises a first step wherein computational means are employed to perform a fitting operation between the chemical entity and a compound according to any one of claims 1-3, or a binding pocket according to any one of claims 4-6, and a second step wherein the results of said fitting operation are analysed to quantify the association between the chemical entity and the compound or binding pocket.
12. A method of identifying a potential ligand to a human κ-Fab constant part-comprising composition, which method comprises

(a) generating a three-dimensional structure of a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6;

(b) employing said three-dimensional structure to design a candidate ligand;

(c) providing said candidate ligand;

(d) contacting the candidate ligand with a human κ-Fab constant part-comprising composition comprising said compound or binding pocket to verify any binding; and, optionally,

(e) repeating steps (b)-(d).

13. A method for evaluating the potential or ability of a chemical entity to associate with a human κ-Fab constant part-comprising composition, which method comprises the steps of

(a) providing a virtual library of chemical entities;

(b) docking the chemical entities to a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6;

(c) defining at least one query based on the results of the docking operation;

(d) screening all entities docked in step (b) while in the docked conformation with the query defined in step (c) for evaluating the potential or ability thereof to associate to the compound or binding pocket;

(e) inspection and, optionally, removal of redundancy; and

(f) providing one or more of the chemical entities that associated with the compound or binding pocket and experimentally testing their binding to a human κ-Fab constant part-comprising composition; and, if more than one chemical entity was tested,

(g) rating the affinities thereof to human κ-Fab constant part-comprising composition.

14. A method according to claim 13, wherein step (a) further comprises a subsequent step of filtering and removal of redundancy among the entities of the library provided.

15. A method according to claim 13 or 14, wherein the results of the docking operation of step (b) are evaluated by visual inspection of the contact between the interacting surface of the compound or binding pocket and the molecular surface(s).

16. Use of a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6 for identification or isolation of a ligand

capable of selective binding of a human κ-Fab constant part-comprising composition.

17. Use of a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6 in site-specific modification of a human κ-Fab constant part-comprising composition.

18. Use according to claim 17, wherein the modification is a stabilisation of Fab-folding by binding a ligand selectively to the compound or binding pocket.

19. Use of a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6 in an immunological assay for detection of a human κ-Fab constant part-comprising composition.

20. A computer for producing a three-dimensional representation of a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6, which computer comprises

(i) a computer-readable data storage medium comprising a data storage material encoded with computer-readable data, wherein said data comprises the structure coordinates as shown in Fig 1 for an IgG κ light chain for the amino acids Q124, S127, G128, T129, S131, V133, G157, N158, S159, Q160, E161, S162, S176, S177, T178, T180, and L181 and the structure coordinates as shown in Fig 1 for an IgG heavy chain for the amino acids P128, S129, L133, L150, K152, F175, P176, V178, L179, Q180, L184, L187 and S188;

(ii) a working memory for storing instructions for processing said computer-readable data;

(iii) a central-processing unit coupled to said working memory and to said computer-readable data storage medium for processing said computer-machine readable data into said three-dimensional representation; and

(iv) a display coupled to said central-processing unit for displaying said three-dimensional representation.

21. A computer according to claim 20, wherein the computer-readable data further comprises the structure coordinates as shown in Fig 1 for an IgG heavy chain for the amino acids K126, F131, D153, S181, S182, and S186.

22. A machine-readable datastorage medium comprising a data storage material encoded with machine-readable data, wherein said data is defined by all or a portion of the structure coordinates of a compound according to any one of claims 1-3 or a binding pocket according to any one of claims 4-6.

ABSTRACT

The present invention relates to a human IgG binding pocket comprised of a first interacting surface, which originates from an IgG κ light chain, and a second interacting surface, which originates from an IgG heavy chain, which amino acids are strictly conserved between human IgGs of κ-type, or a functional derivative of said binding pocket. The invention can also be defined as an isolated compound, which comprises molecules that defines the shape of said binding pocket. Further, the invention relates to various methods of using the novel binding pocket, such as in screening for identification of chemical entities capable of selective binding thereof, and in other experimental and/or virtual methods for design and/or identification of chemical entities capable of selective binding thereof.

Fig 1a kappa light chain

Atom type						X	Y	Z	Occ	B	
ATOM	928	N	GLN	L	124	-44.718	27.024	79.393	1.00	37.64	N
ATOM	929	CA	GLN	L	124	-43.847	25.897	79.535	1.00	38.32	C
ATOM	930	C	GLN	L	124	-44.309	25.088	80.734	1.00	39.17	C
ATOM	931	O	GLN	L	124	-44.458	23.876	80.578	1.00	40.06	O
ATOM	932	CB	GLN	L	124	-42.414	26.311	79.745	1.00	37.76	C
ATOM	933	CG	GLN	L	124	-41.615	25.026	79.581	1.00	34.56	C
ATOM	934	CD	GLN	L	124	-40.133	25.152	79.698	1.00	34.95	C
ATOM	935	OE1	GLN	L	124	-39.440	24.138	79.682	1.00	34.80	O
ATOM	936	NE2	GLN	L	124	-39.569	26.344	79.820	1.00	39.75	N
ATOM	954	N	SER	L	127	-46.898	22.499	80.067	1.00	50.58	N
ATOM	955	CA	SER	L	127	-46.559	21.169	79.588	1.00	49.80	C
ATOM	956	C	SER	L	127	-45.890	20.274	80.637	1.00	49.81	C
ATOM	957	O	SER	L	127	-45.283	19.248	80.318	1.00	50.44	O
ATOM	958	CB	SER	L	127	-45.674	21.333	78.368	1.00	50.26	C
ATOM	959	OG	SER	L	127	-44.618	22.263	78.551	1.00	51.43	O
ATOM	960	N	GLY	L	128	-45.954	20.623	81.919	1.00	48.65	N
ATOM	961	CA	GLY	L	128	-45.371	19.786	82.925	1.00	47.11	C
ATOM	962	C	GLY	L	128	-43.851	19.873	82.985	1.00	46.88	C
ATOM	963	O	GLY	L	128	-43.322	19.013	83.700	1.00	46.88	O
ATOM	964	N	THR	L	129	-43.091	20.805	82.358	1.00	46.66	N
ATOM	965	CA	THR	L	129	-41.625	20.919	82.516	1.00	43.85	C
ATOM	966	C	THR	L	129	-41.246	22.341	82.832	1.00	37.58	C
ATOM	967	O	THR	L	129	-42.031	23.269	82.637	1.00	35.77	O
ATOM	968	CB	THR	L	129	-40.785	20.528	81.250	1.00	48.03	C
ATOM	969	OG1	THR	L	129	-41.566	20.726	80.058	1.00	54.58	O
ATOM	970	CG2	THR	L	129	-40.269	19.111	81.408	1.00	49.23	C
ATOM	976	N	SER	L	131	-37.741	24.856	82.399	1.00	29.71	N
ATOM	977	CA	SER	L	131	-36.337	25.100	82.108	1.00	27.40	C
ATOM	978	C	SER	L	131	-35.958	26.455	82.672	1.00	24.38	C
ATOM	979	O	SER	L	131	-36.663	27.454	82.446	1.00	23.59	O
ATOM	980	CB	SER	L	131	-36.097	25.078	80.593	1.00	29.26	C
ATOM	981	OG	SER	L	131	-36.672	23.985	79.880	1.00	28.43	O
ATOM	989	N	VAL	L	133	-32.859	29.248	82.770	1.00	23.53	N
ATOM	990	CA	VAL	L	133	-31.671	29.552	81.985	1.00	21.58	C
ATOM	991	C	VAL	L	133	-30.829	30.592	82.700	1.00	21.93	C
ATOM	992	O	VAL	L	133	-31.363	31.514	83.297	1.00	22.42	O
ATOM	993	CB	VAL	L	133	-32.042	30.112	80.607	1.00	21.06	C
ATOM	994	CG1	VAL	L	133	-30.831	30.026	79.693	1.00	25.56	C
ATOM	995	CG2	VAL	L	133	-33.149	29.296	79.958	1.00	24.10	C
ATOM	1188	N	GLY	L	157	-26.853	18.788	90.054	1.00	53.00	N
ATOM	1189	CA	GLY	L	157	-26.116	18.154	88.943	1.00	52.14	C
ATOM	1190	C	GLY	L	157	-27.023	17.720	87.749	1.00	51.94	C
ATOM	1191	O	GLY	L	157	-26.809	16.631	87.208	1.00	52.06	O
ATOM	1192	N	ASN	L	158	-28.025	18.503	87.273	1.00	50.46	N
ATOM	1193	CA	ASN	L	158	-28.946	18.183	86.142	1.00	46.46	C
ATOM	1194	C	ASN	L	158	-29.116	19.347	85.106	1.00	44.90	C
ATOM	1195	O	ASN	L	158	-30.222	19.704	84.625	1.00	40.45	O
ATOM	1196	CB	ASN	L	158	-30.312	17.839	86.692	1.00	47.64	C
ATOM	1197	CG	ASN	L	158	-30.916	19.055	87.386	1.00	52.21	C
ATOM	1198	OD1	ASN	L	158	-30.438	19.529	88.430	1.00	47.79	O
ATOM	1199	ND2	ASN	L	158	-31.930	19.646	86.768	1.00	55.19	N
ATOM	1200	N	SER	L	159	-27.980	19.972	84.739	1.00	41.19	N
ATOM	1201	CA	SER	L	159	-27.943	21.083	83.811	1.00	38.82	C
ATOM	1202	C	SER	L	159	-27.136	20.769	82.554	1.00	38.54	C
ATOM	1203	O	SER	L	159	-26.262	19.891	82.610	1.00	39.32	O
ATOM	1204	CB	SER	L	159	-27.323	22.257	84.505	1.00	32.87	C
ATOM	1205	OG	SER	L	159	-26.007	21.915	84.912	1.00	34.12	O
ATOM	1206	N	GLN	L	160	-27.397	21.485	81.451	1.00	37.38	N
ATOM	1207	CA	GLN	L	160	-26.575	21.389	80.266	1.00	35.88	C
ATOM	1208	C	GLN	L	160	-26.118	22.789	79.886	1.00	32.74	C
ATOM	1209	O	GLN	L	160	-26.831	23.765	80.112	1.00	28.45	O
ATOM	1210	CB	GLN	L	160	-27.325	20.798	79.077	1.00	40.64	C
ATOM	1211	CG	GLN	L	160	-27.352	19.273	79.129	1.00	47.64	C
ATOM	1212	CD	GLN	L	160	-27.353	18.619	77.751	1.00	51.42	C
ATOM	1213	OE1	GLN	L	160	-26.474	17.841	77.354	1.00	54.61	O

Fig 1a kappa light chain

ATOM	1214	NE2	GLN	L	160	-28.351	18.941	76.956	1.00	51.87	N
ATOM	1215	N	GLU	L	161	-24.947	22.884	79.252	1.00	32.26	N
ATOM	1216	CA	GLU	L	161	-24.315	24.116	78.812	1.00	30.57	C
ATOM	1217	C	GLU	L	161	-24.096	24.228	77.315	1.00	29.51	C
ATOM	1218	O	GLU	L	161	-24.030	23.218	76.609	1.00	31.47	O
ATOM	1219	CB	GLU	L	161	-22.989	24.254	79.465	1.00	31.63	C
ATOM	1220	CG	GLU	L	161	-23.068	25.232	80.584	1.00	39.52	C
ATOM	1221	CD	GLU	L	161	-22.438	24.715	81.857	1.00	45.11	C
ATOM	1222	OE1	GLU	L	161	-21.196	24.764	81.949	1.00	43.57	O
ATOM	1223	OE2	GLU	L	161	-23.211	24.287	82.736	1.00	48.88	O
ATOM	1224	N	SER	L	162	-23.964	25.449	76.818	1.00	27.95	N
ATOM	1225	CA	SER	L	162	-23.733	25.712	75.415	1.00	24.52	C
ATOM	1226	C	SER	L	162	-22.917	27.003	75.355	1.00	23.12	C
ATOM	1227	O	SER	L	162	-23.213	27.968	76.057	1.00	21.32	O
ATOM	1228	CB	SER	L	162	-25.089	25.831	74.776	1.00	24.91	C
ATOM	1229	OG	SER	L	162	-24.944	26.008	73.380	1.00	28.23	O
ATOM	1332	N	SER	L	176	-24.700	29.533	78.016	1.00	20.73	N
ATOM	1333	CA	SER	L	176	-25.984	29.359	78.650	1.00	20.18	C
ATOM	1334	C	SER	L	176	-25.967	28.050	79.391	1.00	19.90	C
ATOM	1335	O	SER	L	176	-25.400	27.058	78.938	1.00	18.83	O
ATOM	1336	CB	SER	L	176	-27.081	29.343	77.602	1.00	22.81	C
ATOM	1337	OG	SER	L	176	-26.755	28.427	76.557	1.00	27.50	O
ATOM	1338	N	SER	L	177	-26.543	28.045	80.570	1.00	21.10	N
ATOM	1339	CA	SER	L	177	-26.716	26.843	81.325	1.00	22.83	C
ATOM	1340	C	SER	L	177	-28.233	26.701	81.427	1.00	24.50	C
ATOM	1341	O	SER	L	177	-28.927	27.679	81.752	1.00	26.47	O
ATOM	1342	CB	SER	L	177	-26.100	27.030	82.675	1.00	20.36	C
ATOM	1343	OG	SER	L	177	-25.923	25.738	83.209	1.00	25.00	O
ATOM	1344	N	THR	L	178	-28.783	25.535	81.113	1.00	26.21	N
ATOM	1345	CA	THR	L	178	-30.193	25.289	81.284	1.00	25.67	C
ATOM	1346	C	THR	L	178	-30.333	24.182	82.316	1.00	26.52	C
ATOM	1347	O	THR	L	178	-29.692	23.127	82.251	1.00	25.41	O
ATOM	1348	CB	THR	L	178	-30.797	24.854	79.993	1.00	24.43	C
ATOM	1349	OG1	THR	L	178	-30.504	25.890	79.065	1.00	27.73	O
ATOM	1350	CG2	THR	L	178	-32.288	24.606	80.101	1.00	23.92	C
ATOM	1359	N	THR	L	180	-33.064	21.776	83.928	1.00	33.72	N
ATOM	1360	CA	THR	L	180	-34.412	21.334	83.617	1.00	36.96	C
ATOM	1361	C	THR	L	180	-34.895	20.441	84.742	1.00	39.75	C
ATOM	1362	O	THR	L	180	-34.162	19.554	85.220	1.00	40.12	O
ATOM	1363	CB	THR	L	180	-34.439	20.578	82.248	1.00	37.34	C
ATOM	1364	OG1	THR	L	180	-34.262	21.580	81.236	1.00	38.56	O
ATOM	1365	CG2	THR	L	180	-35.746	19.829	81.975	1.00	36.31	C
ATOM	1366	N	LEU	L	181	-36.102	20.772	85.213	1.00	41.45	N
ATOM	1367	CA	LEU	L	181	-36.790	19.955	86.189	1.00	41.68	C
ATOM	1368	C	LEU	L	181	-38.283	19.907	85.844	1.00	41.64	C
ATOM	1369	O	LEU	L	181	-38.823	20.667	85.022	1.00	39.32	O
ATOM	1370	CB	LEU	L	181	-36.472	20.527	87.616	1.00	41.26	C
ATOM	1371	CG	LEU	L	181	-36.887	21.835	88.321	1.00	44.99	C
ATOM	1372	CD1	LEU	L	181	-35.940	21.997	89.487	1.00	42.76	C
ATOM	1373	CD2	LEU	L	181	-36.694	23.093	87.505	1.00	45.40	C

Fig 1b heavy chain

Atom type						X	Y	Z	occ	B	
ATOM	2595	N	LYS	H	126	-39.678	16.046	64.413	1.00	20.92	N
ATOM	2596	CA	LYS	H	126	-40.480	16.635	65.460	1.00	19.79	C
ATOM	2597	C	LYS	H	126	-40.194	18.131	65.371	1.00	21.17	C
ATOM	2598	O	LYS	H	126	-39.025	18.554	65.306	1.00	23.97	O
ATOM	2599	CB	LYS	H	126	-40.054	16.081	66.825	1.00	18.88	C
ATOM	2600	CG	LYS	H	126	-41.090	16.590	67.794	1.00	26.44	C
ATOM	2601	CD	LYS	H	126	-40.944	16.341	69.291	1.00	32.32	C
ATOM	2602	CE	LYS	H	126	-41.916	17.243	70.134	1.00	37.31	C
ATOM	2603	NZ	LYS	H	126	-41.584	18.677	70.172	1.00	35.44	N
ATOM	2608	N	PRO	H	128	-40.310	22.204	66.796	1.00	18.58	N
ATOM	2609	CA	PRO	H	128	-39.950	22.699	68.117	1.00	19.70	C
ATOM	2610	C	PRO	H	128	-41.041	23.367	68.948	1.00	22.10	C
ATOM	2611	O	PRO	H	128	-42.127	23.691	68.475	1.00	25.30	O
ATOM	2612	CB	PRO	H	128	-38.769	23.602	67.812	1.00	18.78	C
ATOM	2613	CG	PRO	H	128	-39.053	24.200	66.457	1.00	17.91	C
ATOM	2614	CD	PRO	H	128	-39.898	23.122	65.749	1.00	20.81	C
ATOM	2615	N	SER	H	129	-40.828	23.620	70.221	1.00	24.82	N
ATOM	2616	CA	SER	H	129	-41.770	24.395	70.995	1.00	23.50	C
ATOM	2617	C	SER	H	129	-40.946	25.623	71.266	1.00	24.48	C
ATOM	2618	O	SER	H	129	-39.763	25.502	71.565	1.00	22.58	O
ATOM	2619	CB	SER	H	129	-42.105	23.686	72.286	1.00	28.31	C
ATOM	2620	OG	SER	H	129	-42.934	22.546	72.073	1.00	36.78	O
ATOM	2628	N	PHE	H	131	-40.521	28.925	73.626	1.00	30.04	N
ATOM	2629	CA	PHE	H	131	-41.040	29.482	74.848	1.00	27.87	C
ATOM	2630	C	PHE	H	131	-40.215	30.723	75.051	1.00	30.92	C
ATOM	2631	O	PHE	H	131	-39.007	30.683	74.789	1.00	26.23	O
ATOM	2632	CB	PHE	H	131	-40.810	28.570	76.022	1.00	30.11	C
ATOM	2633	CG	PHE	H	131	-41.537	27.235	75.970	1.00	30.98	C
ATOM	2634	CD1	PHE	H	131	-42.931	27.183	75.945	1.00	30.57	C
ATOM	2635	CD2	PHE	H	131	-40.808	26.050	75.966	1.00	33.12	C
ATOM	2636	CE1	PHE	H	131	-43.590	25.948	75.915	1.00	31.13	C
ATOM	2637	CE2	PHE	H	131	-41.479	24.815	75.932	1.00	33.57	C
ATOM	2638	CZ	PHE	H	131	-42.863	24.765	75.907	1.00	31.34	C
ATOM	2646	N	LEU	H	133	-38.146	33.716	77.032	1.00	38.18	N
ATOM	2647	CA	LEU	H	133	-37.285	33.930	78.190	1.00	34.16	C
ATOM	2648	C	LEU	H	133	-37.523	35.428	78.330	1.00	35.44	C
ATOM	2649	O	LEU	H	133	-37.005	36.294	77.609	1.00	32.35	O
ATOM	2650	CB	LEU	H	133	-35.823	33.622	77.863	1.00	29.24	C
ATOM	2651	CG	LEU	H	133	-35.533	32.258	77.309	1.00	22.14	C
ATOM	2652	CD1	LEU	H	133	-34.066	32.136	77.012	1.00	23.67	C
ATOM	2653	CD2	LEU	H	133	-35.970	31.213	78.300	1.00	27.77	C
ATOM	2749	N	LEU	H	150	-36.371	30.246	73.846	1.00	22.90	N
ATOM	2750	CA	LEU	H	150	-35.971	28.876	74.075	1.00	23.38	C
ATOM	2751	C	LEU	H	150	-36.705	28.058	73.003	1.00	25.45	C
ATOM	2752	O	LEU	H	150	-37.917	28.204	72.817	1.00	24.96	O
ATOM	2753	CB	LEU	H	150	-36.391	28.505	75.477	1.00	18.99	C
ATOM	2754	CG	LEU	H	150	-36.325	27.052	75.868	1.00	19.75	C
ATOM	2755	CD1	LEU	H	150	-34.917	26.528	75.789	1.00	22.45	C
ATOM	2756	CD2	LEU	H	150	-36.781	26.912	77.286	1.00	19.55	C
ATOM	2764	N	LYS	H	152	-37.287	24.376	72.183	1.00	25.67	N
ATOM	2765	CA	LYS	H	152	-37.209	23.103	72.858	1.00	23.11	C
ATOM	2766	C	LYS	H	152	-37.793	21.909	72.110	1.00	23.19	C
ATOM	2767	O	LYS	H	152	-38.886	21.985	71.563	1.00	22.11	O
ATOM	2768	CB	LYS	H	152	-37.905	23.319	74.200	1.00	25.76	C
ATOM	2769	CG	LYS	H	152	-37.302	22.378	75.195	1.00	29.14	C
ATOM	2770	CD	LYS	H	152	-37.759	22.579	76.622	1.00	30.74	C
ATOM	2771	CE	LYS	H	152	-36.922	21.597	77.460	1.00	28.69	C
ATOM	2772	NZ	LYS	H	152	-37.314	20.228	77.199	1.00	25.73	N
ATOM	2773	N	ASP	H	153	-37.045	20.807	72.047	1.00	25.81	N
ATOM	2774	CA	ASP	H	153	-37.461	19.487	71.575	1.00	22.60	C
ATOM	2775	C	ASP	H	153	-37.870	19.231	70.146	1.00	20.15	C
ATOM	2776	O	ASP	H	153	-38.939	18.761	69.803	1.00	18.56	O
ATOM	2777	CB	ASP	H	153	-38.561	19.010	72.523	1.00	26.65	C
ATOM	2778	CG	ASP	H	153	-38.083	18.807	73.962	1.00	26.68	C
ATOM	2779	OD1	ASP	H	153	-36.935	18.446	74.194	1.00	28.52	O

Fig 1b heavy chain

ATOM	2780	OD2	ASP	H	153	-38.866	19.018	74.873	1.00	26.88	O
ATOM	2940	N	PHE	H	175	-27.214	30.210	70.335	1.00	23.94	N
ATOM	2941	CA	PHE	H	175	-26.383	29.122	70.813	1.00	23.42	C
ATOM	2942	C	PHE	H	175	-26.478	27.831	69.986	1.00	23.74	C
ATOM	2943	O	PHE	H	175	-27.538	27.522	69.409	1.00	23.81	O
ATOM	2944	CB	PHE	H	175	-26.758	28.815	72.248	1.00	22.94	C
ATOM	2945	CG	PHE	H	175	-26.259	29.899	73.148	1.00	20.21	C
ATOM	2946	CD1	PHE	H	175	-24.971	29.801	73.645	1.00	19.49	C
ATOM	2947	CD2	PHE	H	175	-27.079	30.977	73.458	1.00	20.84	C
ATOM	2948	CE1	PHE	H	175	-24.497	30.807	74.468	1.00	20.36	C
ATOM	2949	CE2	PHE	H	175	-26.595	31.980	74.294	1.00	22.58	C
ATOM	2950	CZ	PHE	H	175	-25.300	31.901	74.800	1.00	21.02	C
ATOM	2951	N	PRO	H	176	-25.360	27.078	69.878	1.00	22.56	N
ATOM	2952	CA	PRO	H	176	-25.321	25.723	69.318	1.00	19.83	C
ATOM	2953	C	PRO	H	176	-26.377	24.835	69.977	1.00	21.20	C
ATOM	2954	O	PRO	H	176	-26.508	24.942	71.200	1.00	22.55	O
ATOM	2955	CB	PRO	H	176	-23.910	25.305	69.595	1.00	16.67	C
ATOM	2956	CG	PRO	H	176	-23.083	26.559	69.637	1.00	15.02	C
ATOM	2957	CD	PRO	H	176	-24.018	27.503	70.334	1.00	17.20	C
ATOM	2963	N	VAL	H	178	-28.150	21.582	71.822	1.00	22.66	N
ATOM	2964	CA	VAL	H	178	-27.623	20.460	72.565	1.00	21.24	C
ATOM	2965	C	VAL	H	178	-28.654	19.371	72.365	1.00	20.88	C
ATOM	2966	O	VAL	H	178	-29.868	19.553	72.269	1.00	22.54	O
ATOM	2967	CB	VAL	H	178	-27.441	20.749	74.109	1.00	23.34	C
ATOM	2968	CG1	VAL	H	178	-26.426	21.863	74.326	1.00	21.50	C
ATOM	2969	CG2	VAL	H	178	-28.744	21.171	74.737	1.00	25.02	C
ATOM	2970	N	LEU	H	179	-28.110	18.208	72.193	1.00	22.43	N
ATOM	2971	CA	LEU	H	179	-28.876	17.011	72.085	1.00	25.70	C
ATOM	2972	C	LEU	H	179	-29.097	16.527	73.522	1.00	25.97	C
ATOM	2973	O	LEU	H	179	-28.187	16.399	74.348	1.00	25.39	O
ATOM	2974	CB	LEU	H	179	-28.076	16.026	71.278	1.00	25.57	C
ATOM	2975	CG	LEU	H	179	-28.702	14.674	71.023	1.00	27.43	C
ATOM	2976	CD1	LEU	H	179	-29.897	14.757	70.074	1.00	19.28	C
ATOM	2977	CD2	LEU	H	179	-27.587	13.805	70.469	1.00	30.31	C
ATOM	2978	N	GLN	H	180	-30.365	16.320	73.815	1.00	27.28	N
ATOM	2979	CA	GLN	H	180	-30.821	15.886	75.111	1.00	25.86	C
ATOM	2980	C	GLN	H	180	-30.787	14.360	75.199	1.00	26.76	C
ATOM	2981	O	GLN	H	180	-30.630	13.675	74.180	1.00	27.19	O
ATOM	2982	CB	GLN	H	180	-32.233	16.463	75.292	1.00	28.23	C
ATOM	2983	CG	GLN	H	180	-32.316	17.984	75.105	1.00	28.44	C
ATOM	2984	CD	GLN	H	180	-33.725	18.562	75.115	1.00	31.65	C
ATOM	2985	OE1	GLN	H	180	-34.406	18.608	74.093	1.00	30.70	O
ATOM	2986	NE2	GLN	H	180	-34.230	19.012	76.261	1.00	30.98	N
ATOM	2987	N	SER	H	181	-30.940	13.753	76.391	1.00	28.39	N
ATOM	2988	CA	SER	H	181	-30.945	12.305	76.549	1.00	28.80	C
ATOM	2989	C	SER	H	181	-32.113	11.663	75.787	1.00	25.40	C
ATOM	2990	O	SER	H	181	-31.965	10.542	75.300	1.00	28.76	O
ATOM	2991	CB	SER	H	181	-30.979	12.001	78.067	1.00	31.94	C
ATOM	2992	OG	SER	H	181	-31.812	12.915	78.815	1.00	40.94	O
ATOM	2993	N	SER	H	182	-33.258	12.324	75.579	1.00	21.90	N
ATOM	2994	CA	SER	H	182	-34.325	11.787	74.720	1.00	24.38	C
ATOM	2995	C	SER	H	182	-33.959	11.687	73.227	1.00	25.28	C
ATOM	2996	O	SER	H	182	-34.562	10.902	72.497	1.00	29.85	O
ATOM	2997	CB	SER	H	182	-35.556	12.654	74.850	1.00	17.40	C
ATOM	2998	OG	SER	H	182	-35.104	13.995	74.772	1.00	19.22	O
ATOM	3003	N	LEU	H	184	-33.775	14.556	71.267	1.00	20.48	N
ATOM	3004	CA	LEU	H	184	-34.278	15.749	70.637	1.00	17.63	C
ATOM	3005	C	LEU	H	184	-33.314	16.869	71.000	1.00	18.68	C
ATOM	3006	O	LEU	H	184	-32.549	16.765	71.956	1.00	16.48	O
ATOM	3007	CB	LEU	H	184	-35.675	15.980	71.168	1.00	18.23	C
ATOM	3008	CG	LEU	H	184	-36.724	14.864	71.080	1.00	12.53	C
ATOM	3009	CD1	LEU	H	184	-37.909	15.249	71.922	1.00	10.58	C
ATOM	3010	CD2	LEU	H	184	-37.141	14.621	69.658	1.00	13.49	C
ATOM	3023	N	SER	H	186	-32.310	21.176	71.626	1.00	19.45	N
ATOM	3024	CA	SER	H	186	-32.755	22.411	72.223	1.00	20.48	C
ATOM	3025	C	SER	H	186	-31.701	23.450	71.937	1.00	23.05	C
ATOM	3026	O	SER	H	186	-30.521	23.102	71.874	1.00	25.18	O
ATOM	3027	CB	SER	H	186	-32.916	22.306	73.718	1.00	21.58	C

Fig 1b heavy chain

ATOM	3028	OG	SER	H	186	-34.253	21.920	74.021	1.00	32.73	O
ATOM	3029	N	LEU	H	187	-32.104	24.707	71.768	1.00	21.56	N
ATOM	3030	CA	LEU	H	187	-31.233	25.811	71.415	1.00	21.58	C
ATOM	3031	C	LEU	H	187	-31.765	27.082	72.120	1.00	23.47	C
ATOM	3032	O	LEU	H	187	-32.948	27.118	72.496	1.00	24.42	O
ATOM	3033	CB	LEU	H	187	-31.309	25.838	69.897	1.00	19.86	C
ATOM	3034	CG	LEU	H	187	-30.875	26.971	69.054	1.00	21.75	C
ATOM	3035	CD1	LEU	H	187	-30.413	26.485	67.691	1.00	19.38	C
ATOM	3036	CD2	LEU	H	187	-32.048	27.868	68.864	1.00	23.32	C
ATOM	3037	N	SER	H	188	-31.014	28.142	72.424	1.00	22.73	N
ATOM	3038	CA	SER	H	188	-31.587	29.401	72.873	1.00	21.20	C
ATOM	3039	C	SER	H	188	-31.069	30.509	71.988	1.00	20.80	C
ATOM	3040	O	SER	H	188	-29.961	30.400	71.441	1.00	21.00	O
ATOM	3041	CB	SER	H	188	-31.179	29.775	74.274	1.00	25.10	C
ATOM	3042	OG	SER	H	188	-31.586	28.721	75.127	1.00	31.30	O

Fig 2(a)

Human Fab constant light kappa alignment

```
                         1              1               1               1               1               1
                         0              1               2               3               4               5
                    34567890123456789012345678901234567890123456789012345678901234567890123456789012
S02577_Ig_Kappa     ----------------VAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
AAB50880_IgG1_Kappa T-PLTFGQGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVBLLNNFYPREAKVQWKVDN
A23746_Ig_Kappa     S-PLTFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVGLLNNFYPREAKVQWKVDN
AAA58921_Ig_Kappa   ---------------TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYHREAKVQWKVDN
AAB86466_Ig_Kappa   W-FRTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQRKVDN
AAA58922_Ig_Kappa   ---------------TVAAPSVFIFPPSDEQLKSGTASVVCLLDDFYPREAKVQWKVDN
223335_Ig_Kappa     F-PYTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
CAA09181_Ig_Kappa   F-PFTFGPGTKVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
CAA09183_Ig_Kappa   YSPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
                                    *********************** **;:** ****** ****

                         1              1               1               1               2               2
                         6              7               8               9               0               1
                    3456789012345678901234567890123456789012345678901234567890123456789012345678901234
S02577_Ig_Kappa     ALQSGNSQESVTZZDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
AAB50880_IgG1_Kappa ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYABEVTHQGLSSPVTKSFNRGEB
A23746_Ig_Kappa     ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEG
AAA58921_Ig_Kappa   ALQSGNSQESVTEQESKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEC
AAB86466_Ig_Kappa   ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
AAA58922_Ig_Kappa   ALQSGNSQESVTEQESKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
223335_Ig_Kappa     ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
CAA09181_Ig_Kappa   ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
CAA09183_Ig_Kappa   ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE---------------------------
                    ************ ;*********************
```

Fig. 2(b)

Human Fab, heavy chain constant part alignment

```
                                1         1         1         1         1
                     9          0         1         2         3         4
                     901234567890123456789012345678901234567890123456789012345678
AAH06402_igg3        IELDIV---VVPAPMRGSLGFDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
GC1_igg1             --------------------------------ASTKGPSVFPLAPSSKSTSGGTAALG
A23511_igg3          --------------------------------ASTKGPSVFPLAPCSRSTSGGTAALG
CAC20456_igg3        --------------------------------ASFKGPSVFPLAPCSRSTPGGTAALG
CAA67886_igg3        ---------------------------------STKGPSVFPLAPCSRSTSGGTAALG
AAG00911_igg3        ---------------------------------FTKGPSVFPLAPCSRSTSGGTAALG
AAB59393_igg2        ---------------------------------STKGPSVFPLAPCSRSTSESTAALG
AAG00910_igg2        ----------IIYFDYADFIMDYWGQGTTVTVSTASTKGPSVFPLAPCSRSTSESTAALG
GC2_igg2             --------------------------------ASTKGPSVFPLAPCSRSTSESTAALG
CAC20455_igg2        --------------------------------ASFKGPSVFPLAPCSRSTSESTAALG
AAB59394_igg4        ---------------------------------STKGPSVFPLAPCSRSTSESTAALG
AAG00912_igg4        ----------IIYFDYADFIMDYWGQGTTVTVSTASTKGPSVFPLAPCSRSTSESTAALG
CAC20457_igg4        --------------------------------ASFKGPSVFPLVPCSRSTSESTAALG
AAG00909_igg1        ---------------------------------STKGPSVFPLAPSSKSTSGGTAALG
229601_igg1          RDTAMF---FAHWG-----------QGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
2ig2_H_igg1          GGHGFC---SS----ASCFGPDYWGQGTPVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
1mco_H_igg1          ----------------VPLVVNPWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AAB28159_igg1K       --------------------FAYWGQGTLVTVSAAKTTPPSVYPLAPGSAAQTNSMVTLG
1aqk_H_igg           ---------VLFQQLVLYAPFDIWGQGTMVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
8fab_H_igg1          --------RDYYDSGGYFTVAFDIWGQGTMVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AAA70227_igg1        GAGVTLVRGAIKPSPDYYYGMDVWGQGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
2fb4_H_igg           G-------GHGFCSSASCFGPDYWGQGTPVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AAB50879_igg1        G----------------DVYNRQWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AF184764_igg1        G----------------DVYNRQWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AF184765_igg2        G----------------DVYNRQWGQGTLVTVSSASTKGPSVFPLAPCSRSTSESTAALG
1vge_H_igg1K         P------------YGGGKSEFDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
A49444_igg1          LIAGG--------------IDVWGQGSLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AAB24269_igg1        LIAGG--------------IDVWGQGSLVTVSSASTKGPSVFPLAPSSKSTSGGTAALG
AAA70228_igg3        ETMASR-----------KRAFDIWGQGTMVTVSAASTKGPSVFPLAPCSRSTSGGTAALG
                                                 . ***:**.* * : . . .:**
```

Fig. 2(b)

```
                                 1         1         1         1         1         2
                                 5         6         7         8         9         0
                        901234567890123456789012345678901234567890123456789012345678
AAH06402_igg3           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
GC1_igg1                CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
A23511_igg3             CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVN
CAC20456_igg3           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVN
CAA67886_igg3           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVN
AAG00911_igg3           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVN
AAB59393_igg2           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTCNVD
AAG00910_igg2           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTCNVD
GC2_igg2                CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTCNVD
CAC20455_igg2           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTCNVD
AAB59394_igg4           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQXSGLYSLSSVVTVPSSSLGTKTYTCNVD
AAG00912_igg4           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTCNVD
CAC20457_igg4           CLVKDYFPEPVTVSWNSCALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
AAG00909_igg1           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
229601_igg1             CLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
2ig2_H_igg1             CLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
1mco_H_igg1             CLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQS-DLYTLSSSVTVPSSTWPSETVTCNVA
AAB28159_igg1K          CLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQS-DLYTLSSSVTVPSSTWPSETVTCNVA
1aqk_H_igg              CLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
8fab_H_igg1             CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
AAA70227_igg1           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
2fb4_H_igg              CLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
AAB50879_igg1           BLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIBNVN
AF184764_igg1           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
AF184765_igg2           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTCNVD
1vge_H_igg1K            CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
A49444_igg1             CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVN
AAB24269_igg1           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVSVPSSNLGTQTYTCNVN
AAA70228_igg3           CLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVSVPSSNLGTQTYTCNVN
                         ***.***:****:*** :*:*********** .**:*** *:****.  ::*   **
```

Fig. 2(b)

```
                                  2         2         2         2         2         2
                                  1         2         3         4         5         6
                        901234567890123456789012345678901234567890123456789012345678
AAH06402_igg3           HKPSNTKVDKKVEP--------------------------------------------------
GC1_igg1                HKPSNTKVDKKVEP--------------------------------------------------
A23511_igg3             HKPSNTKVDKRVEL------KTPLGDTTHTCPRCPEPKSCDTPPPCPRCPEPKSCDTPPP
CAC20456_igg3           HKPSNTKVDKRVEL------KTPLGDTTHTCPRCPEPKSCDTPPPCPRCPEPKSCDTPPP
CAA67886_igg3           HKPSNTKVDKRVEL------KTPLGDTTHTCPRCPEPK------------------------
AAG00911_igg3           HKPSNTKVDKRVEL------KTPLGDTPPPCPRCPEPK------------------------
AAB59393_igg2           HKPSNTKVDKTVER------KCCV--------------------------------------
AAG00910_igg2           HKPSNTKVDKTVER------KCCV--------------------------------------
GC2_igg2                HKPSNTKVDKTVER------KCCV--------------------------------------
CAC20455_igg2           HKPSNTKVDKTVER------KCCV--------------------------------------
AAB59394_igg4           HKPSNTKVDKRVES------KYGP--------------------------------------
AAG00912_igg4           HKPSNTKVDKRVES------KYGP--------------------------------------
CAC20457_igg4           HKPSNTKVDKRVES------KYGP--------------------------------------
AAG00909_igg1           HKPSNTKVDKRVEPKSCDKT--------------------------------------------
229601_igg1             HKPSNTKVDKKVEP--------------------------------------------------
2ig2_H_igg1             HKPSNTKVDKRVEP--------------------------------------------------
1mco_H_igg1             HKPSNTKVDKRVAP--------------------------------------------------
AAB28159_igg1K          HPASSTKVDKKIVP--------------------------------------------------
1aqk_H_igg              HKPSNTKVDKKVEP--------------------------------------------------
8fab_H_igg1             HKPSNTKVDKKVEP--------------------------------------------------
AAA70227_igg1           HKPSNTKVDKRVEP--------------------------------------------------
2fb4_H_igg              HKPSNTKVDKRVEP--------------------------------------------------
AAB50879_igg1           HKPSNTK---------------------------------------------------------
AF184764_igg1           HKPSNTKVDKRVEP--------------------------------------------------
AF184765_igg2           HKPSNTKVDKTVER--------------------------------------------------
1vge_H_igg1K            HKPSNTKVDKKVEP--------------------------------------------------
A49444_igg1             HKPSNTKVDKKVEP--------------------------------------------------
AAB24269_igg1           HKPSNTKVDKKVEP--------------------------------------------------
AAA70228_igg3           HKPSNTKVDKTVEL--------------------------------------------------
                        * .*.**
```

Fig. 3